All-sky search of continuous gravitational wave signals

Mikayla Streetman
The University of Texas at Arlington

Mentor:Pia Astone
Università di Roma: La Sapienza
(Dated: August 10, 2016)

The O1 run of LIGO collected data with 4 times the sensitivity of any other run in the history of LIGO. In this paper, we present an analysis of a way to further the final sensitivity with an optimised use of LIGO science times. In our search for continuous waves, we utilize these science times as part of the method for finding a compromise between the total observation time and the quality of the noise. We will discuss how these times affected the analysis and what impact they have on the final sensitivity.

I. INTRODUCTION

In 1916, Albert Einstein predicted the existence of gravitational waves, ripples in space-time caused by energetic processes in the universe, in his theory of general relativity. Examples of these energetic processes are colliding black holes, the merging of neutron stars or white dwarfs, the minimally shaky spin of neutron stars, or many other processes. 100 years later, on 14 September 2016, Einstein's predictions were confirmed when the LIGO detectors observed two black holes merging.

There are four possible classes of sources of gravitational waves in the data, a burst, compact binaries coalescing, the stochastic gravitational wave background, and from an object that emits continuous waves (CW). In this paper, we will be focusing on continuous waves. CW are emitted from isolated NS or from binary systems that are moderately constant and have a precise frequency. They are assumed to produce weak gravitational waves since they are emitted for long periods of time and change very slowly over a long duration. There might be transient CW signals, where the signal lasts for hours or days however, the work we will be focusing on is in regards to the search from isolated NS emitted continuously during the run time. CW emitted by asymmetric rotating neutron stars are among the sources currently searched in the data of interferometric gravitational wave detectors. Roughly 10^9 neutron stars are expected to exist in the Galaxy. Of these, only about 2,400 have been detected through their electromagnetic emission, like pulsars. A fraction of the unseen population of neutron stars could in principle emit gravitational waves in the sensitivity band of detectors, and it is therefore very important to develop efficient data analysis strategies to search the signals they emit.

In general, there are three types of CW searches for isolated NS, emitted continuously during the run. The first, targeted, is when everything about the source is known. The second, directed, is when the only known parameter is the sky position. And third, all-sky, is when all parameters are unknown. In all-sky searches, every combination of parameters must be analyzed until a potential source has been found. Hierarchical methods have been developed to handle the computational problems inherent in all-sky searches.

II. HIERARCHICAL METHOD

A completely coherent method cannot support all-sky searches, because of the copious parameters that pose computational challenges. A completely coherent search, when faced with unpredictable phase variations of the signal during the observation time, would not hold strong. For these reasons, hierarchical schemes have been developed. The first step consists in constructing a short FFT database (SFDB) where each FFT is built from a data chunk of duration, called coherence time, in this analysis they are roughly 8000s. The FFT duration is then a function of the search frequency, with longer FFTs allowed at lower frequencies. From the SFDB we create a time-frequency map, called a peakmap. The peakmap is the input of the incoherent step, based on the FH transform. The most significant candidates are selected at this stage using a coarse grid in the parameter space and an effective way to avoid blinding by particularly disturbed frequency bands. For each coarse candidate a refined search is run again on the neighborhood of the candidate parameters and the final first-level refined candidates are selected. Candidates are then clustered, grouping together those occupying nearby points in the parameter space. In order to significantly reduce the false alarm probability, coincidences are done among clusters of candidates obtained in the analysis of different data sets (of the same detector or of different detectors). Over coincident candidates, after a verification step, a follow-up, with a longer coherence time, is applied. We expect that the use of the FH transform, of the refinement only around coarse candidates, and various aggressive cleaning steps allows to significantly improve the detection efficiency and to partially compensate the shorter FFT length against a lower computational load. [1]

III. DATA QUALITY

The final sensitivity depends on two things, the quality of the noise in the frequency band of the search and the total observation time. The quality of the noise is determined by how much of the noise is considered artificially high or artificially low. The total observation time is the amount of time that data was actually being collected. This is different from the total run time because the data collection could have been paused in order to run tests or make changes to the equipment. In order to obtain the best final sensitivity, we need to find a compromise between the quality of the noise and the total observation time. This is done by keeping the most observation time as possible while also excluding data that is artificially high or low.

IV. ANALYSIS PROCEDURE

The data used was from the first Observing Run (O1) of LIGO. The run was from 15 September 2015 to 16 January 2016. The run came after many improvements to the detectors sensitivity, which increased by 4 times.[2] While this data has been improved by such a great amount, there are still instances of artificially high and low data present. Because of this, it is a very important task is to clean the data even further. Cleaning procedures are used when constructing the frequency data bases, which is the basis for the CW analysis. Short disturbances (glitches) in time are removed, persistent and strong lines in frequency are also removed. And, from the analysis of the spectral data base, weights to be used to further reduce the contribution of disturbed data, are constructed.

A. Science Times

In order to better clean the data, we typically use science times, by removing from the analysis, or putting to zero, all the data which are not part of a science segment. Science times are times in which each detector was collecting data, the data collection was good, and no signal injections were made. They are suspected to be where it would be most likely to find a signal. Each detector separately decides on science times. The science times make up 53.65% of the total observation time for Livingston and 61.78% for Hanford.

B. Spectra Cleaning

We have 4 spectral databases, consisting of Fourier Transforms, of different resolution. The first one, which is what we have analysed, is for searches from 10 to 256 Hz and is what we will focus on. The length of each FFT in this database is 8192 seconds. To use the science times, we created a simple filing system to exclude any data that was not inside science times. However, this was a problem. Since the science times only accounted for just over half of the observation time, the final sensitivity suffers despite containing quality noise. In order to obtain the best final sensitivity we created an exception. This exception is a percentage which allows for more FFT's to be included in the analysis. The percentage essentially says as long as a certain percentage is inside the science time it will be allowed in the analysis. It was decided, in the early stages of the project, that the best percentage to use would be 60% since more than half of the FFT would be inside the science times but it would still leave a small amount of it outside. A further study into the percentages will be done in section V. This method was first applied to a study of spectra, in order to better understand how to use it in the analysis. The goal was to understand how the spectra improved in order to better use them in the adaptive Hough Transform implementation.

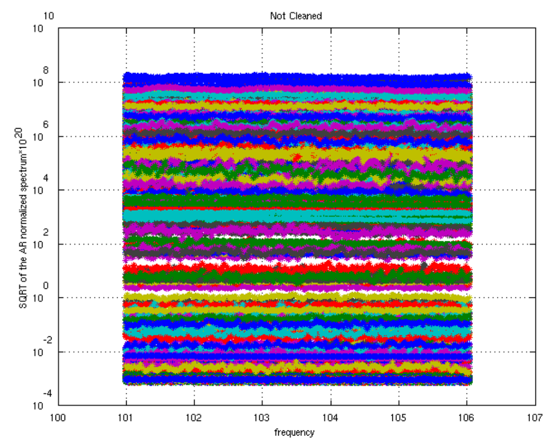

FIG. 1. Graph of 101-106Hz Band with No Veto

Figure 1 is a prime example of the results of using all of the data, in this case for 101-106 Hz. This is called the auto-regressive (AR) spectrum, the average noise of the detector across all frequencies, which is complicated in construction. Essentially it is constructed so that noise estimations are not affected by peaks in frequency and still has the ability to comprehend slow and sudden changes in noise. The spectral amplitude times 10^{20} 1/sqrt(Hz) ranges from 10^{17} to just over 10^{28} and is consistent through all of the frequencies shown. In this situation, artificially low data is not seen but high data is. The average value which we expect at these frequencies is of the order of $5 * 10^{-24}$ to 10^{-23} /sqrt(Hz) so anything above this is considered high. The need for the further cleaning is apparent because of the amount of noise above the average, and for other cases, lower than the average. In Section V, we will discuss the results and the affects our analysis had on this plot.

C. Hough

This same method was applied to the Hough Transform which uses the above spectra to weight the results from each time segment of length 8192 s in this case. The goal of this was to better the critical ratio for each signal thus, making them easier to find. In order to first test the effectiveness of our analysis, we used injected signals with known parameters to enable us to be as precise as possible. Figures 2 through 5 are of injected signals with no veto applied. The plots are estimated spin down vs source frequency. The colored bar is the Hough amplitude, which is proportional to the number of FFTs where the signal (or noise) was present. These plots give us a baseline of Signal to Noise Ratios (SNR) for each injection. In the case of No Science veto, the Hough amplitude is exactly the number of FFTs where signal or noise was present. In the case of the adaptive implementation, used when the veto was applied, the two are proportional. After applying a veto with varying percentages, we will be able to compare the results to these original plots. Results will be discussed in section V.

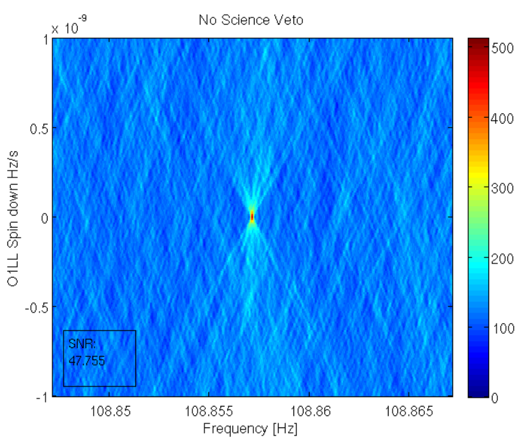

FIG. 2. This is a plot of Pulsar 3, an injected signal, in the Livingston data which has a signal to noise ratio of 47.8 without a veto.

V. RESULTS

A. Spectra

Before the veto, in figure 1, there was a significant amount of noisy data that could be reduced or lowered. After the veto, shown in figure 6, it can be seen that a good amount of noise from 10^6 /sqrt(Hz) to 10^8 /sqrt(Hz) has been removed as well as some noise that is spread throughout the spectrum. It should be noted that artificially high or low data that is still present after the veto is given a lower weight, meaning that it is weighted less than the average data. This is the so called Adaptive implementation of the Hough Transform.

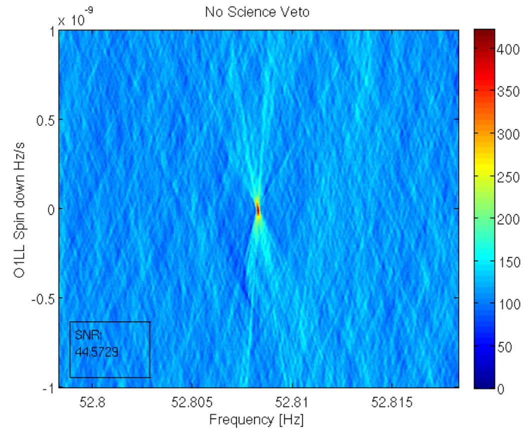

FIG. 3. This is a plot of Pulsar 5, an injected signal, in the Livingston data which has a signal to noise ratio of 44.6 without a veto.

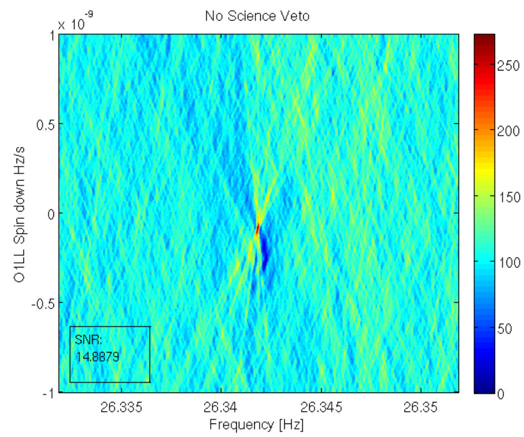

FIG. 4. This is a plot of Pulsar 10, an injected signal, in the Livingston data which has a signal to noise ratio of 14.9 without a veto.

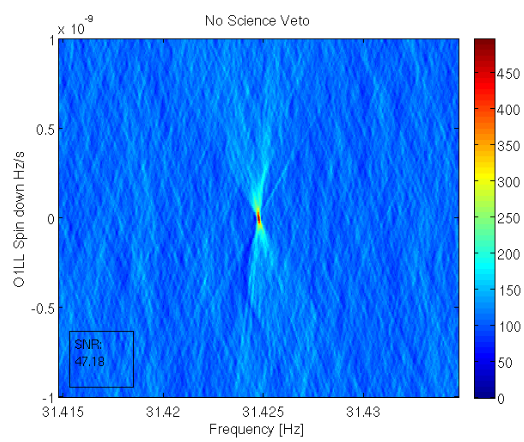

FIG. 5. This is a plot of Pulsar 11, an injected signal, in the Livingston data which has a signal to noise ratio of 47.2 without a veto.

Figure 7, shows a comparison of using and not using a veto. The blue represents the analysis of data without a veto and the red represents the analysis of data with the veto. As you can see, there is a significant difference between the two. When viewed close up, the spectral amplitude, although reduced, follows the same path when cleaned with the veto. This is important to understand because it shows that the veto is not changing the data or modifying it in any way but instead reducing it to a more reasonable level. The improvement in this graph is noticeable, however, it is still not perfect. The green represents a further cleaning done where the data with the percentage veto is cleaned even further by removing data that is 3 times the median.

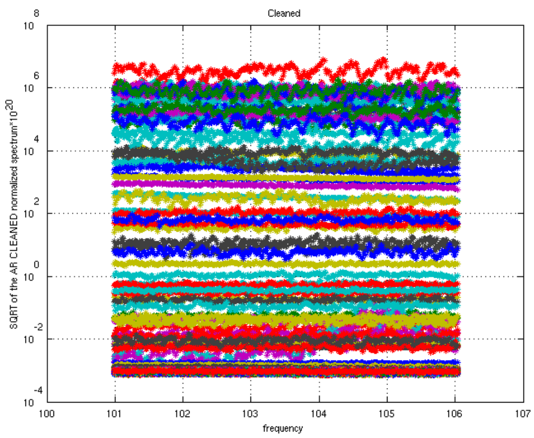

FIG. 6. This is a graph of 101-106Hz Band with a veto of 40%, meaning that 60% of the FFT had to be inside of a science time.

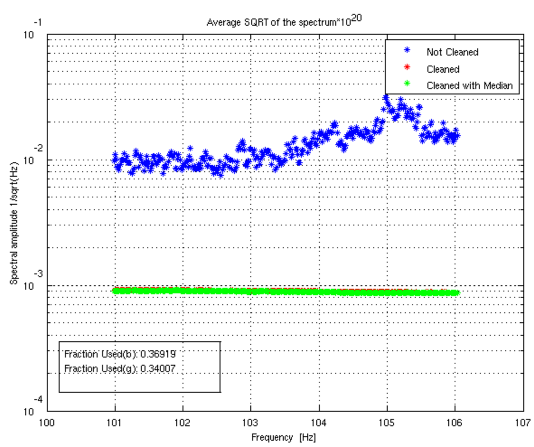

FIG. 7. This graph is a comparison of using all of the data with no veto, using the percentage veto we created, and a further cleaning done by subtracting noise 3 times the median from the data with the percentage veto for the 101-106 Hz band.

B. Hough

In figures 2 through 5 it is difficult to see where improvements could occur. However, compared with figures 8 through 11,which have a veto of 60%, it is easier to see the differences. Besides the distinct change in background color for most all of the figures, the SNR has also improved for each plot. A higher SNR indicates that the final sensitivity has improved, which was our goal.

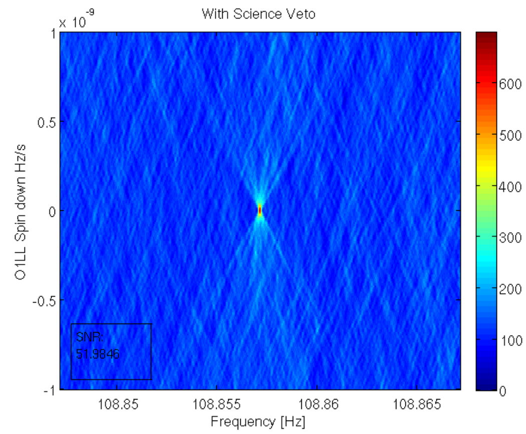

FIG. 8. This is a plot of Pulsar 3, an injected signal, in the Livingston data which has a signal to noise ratio of 59.0 with a veto of 60%

.

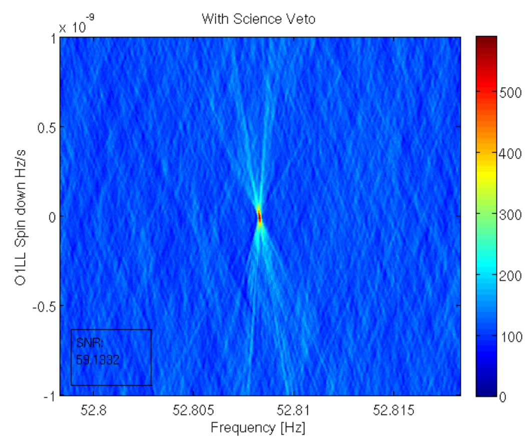

FIG. 9. This is a plot of Pulsar 5, an injected signal, in the Livingston data which has a signal to noise ratio of 59.1 with a veto of 60%

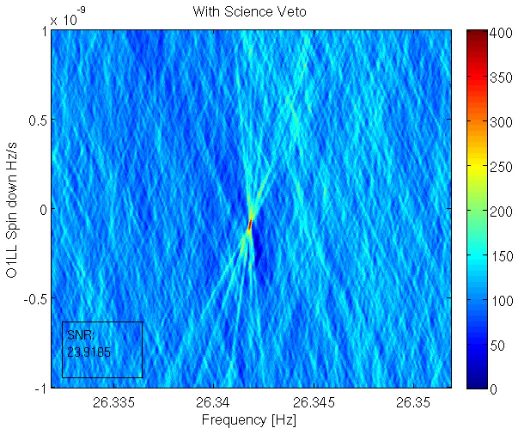

FIG. 10. This is a plot of Pulsar 10, an injected signal, in the Livingston data which has a signal to noise ratio of 23.9 with a veto of 60%.

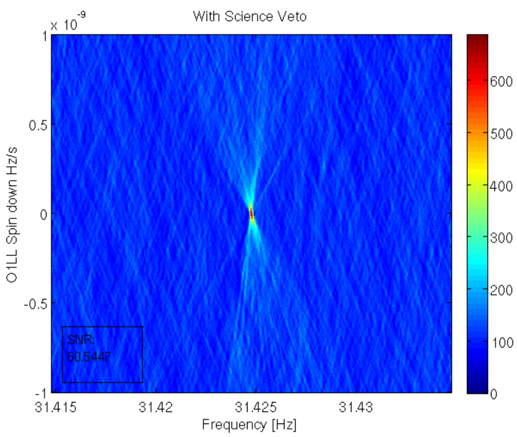

FIG. 11. This is a plot of Pulsar 11, an injected signal, in the Livingston data which has a signal to noise ratio of 60.5 with a veto of 60%.

C. Percentiles

As we ran the analysis with percentiles ranging from 0 to 90, our original thought that 60% would yield the best result was not fully supported by our findings. In order to fully understand which percentage would yield the best results, we created a system that ranked the percentages for each injection and then totalled up how many times a single percentage had the highest or second highest critical ratio for 45 injections. Figures 12 and 13 show 2 of the bands where signals were injected and how they were ranked. In all of 10 the injections shown (five in each band), any veto is better than not using a veto at all, since no veto yields a lower critical ratio than any of the percentages. These two plots give good indications of the trend that the percentages fall into. 10% and 20% usually yield a higher critical ratio than 80% or 90%.

Figure 14 is a tally of how many times a percentage had the highest or second highest critical ratio. So essentially, if 60% yielded the highest critical ratio for an injection then one would be added to the 0.6 column on the graph. This allows us to understand which percentages provided the highest critical ratios. The trend on this plot follows the previous figures where 10% and 20% have much higher critical ratios than higher percentages. Now that we had this information, we wanted to know by how much the critical ratios were improving and this is shown in Figure 15. This plot takes the highest critical ratio for each injection and divides it by the critical ratio obtained by no veto. It can be seen that most of the critical ratios see an improvement of 1.1 to 1.2 times with the percentages. There were only 2 of the 45 injections that did not show improvements and it is possible that they were never recovered and this is just reading noise. But in either scenario, using the veto we created will yield higher critical ratios, more improved data, and positively impact data analysis when searching for continuous waves.

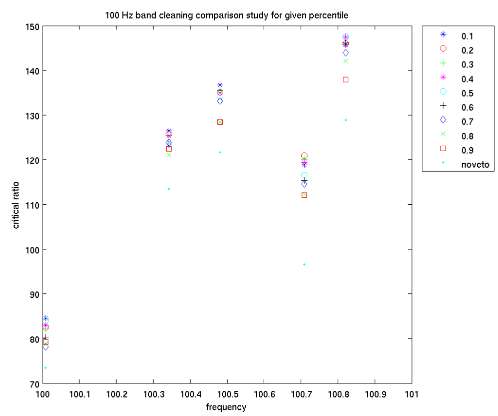

FIG. 12. This is a graph that ranks the percentages, for each of the 5 injections in the 100 Hz band, according to critical ratio.

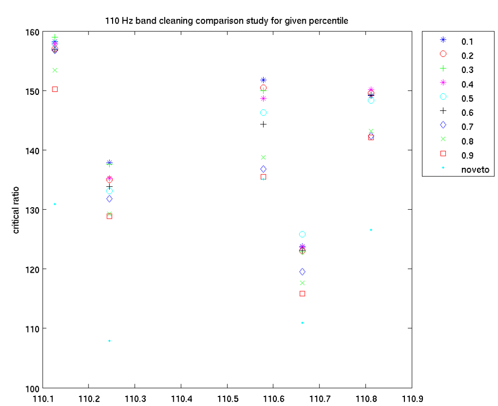

FIG. 13. This is a graph that ranks the percentages, for each of the 5 injections in the 110 Hz band, according to critical ratio.

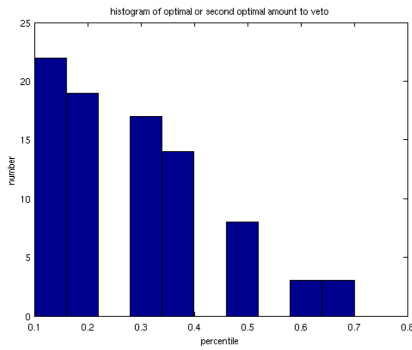

FIG. 14. This is a graph that tallied how many times a percentage was ranked with the highest or second highest critical ratio for all 45 injections.

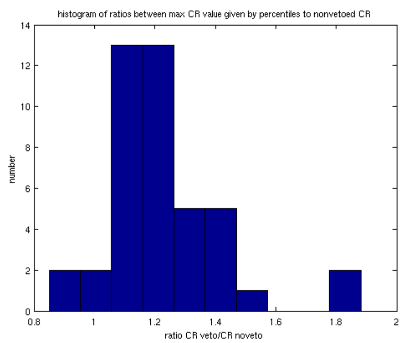

FIG. 15. This is a graph that shows by how much the critical ratios are improved using the percentage that ranked the highest.

VI. CONCLUSIONS

The LIGO data from the O1 run is by far the most sensitive data to date, however, there is always room for improvement. Our goal was to further improve the quality of the sensitivity by finding a compromise between the quality of the noise and the amount of observation time used in the data analysis. By using science times provided by each detector and a percentage that allows for more data to be analyzed, we were able to successfully improve the sensitivity of the search on these data. We determined that any veto is better than not using one at all. By using a percentage of 10% to 20%, an improvement of 1.1 to 1.2 times in signal to noise ratio can be expected in most cases. Signals are easier to find and the final sensitivity has improved. Although our improvement is significant, it is still not where it could be. There is still much more that can be done in order to see the best improvement possible.

VII. ACKNOWLEDGEMENTS

We would like to thank the National Science Foundation for providing the funding that allowed for this research to happen and the University of Florida for orchestrating the program. We would also like to thank Andrew Miller for his work on the injections. And of course, a huge thanks to Pia for being such an amazing mentor and host.

[1] Pia Astone et al., *Method for all-sky searches of continuous gravitational wave signals using the frequency-Hough transform* 2014:

[2] Veronica Kondrashov *First Observing Run Ends.* 2016: